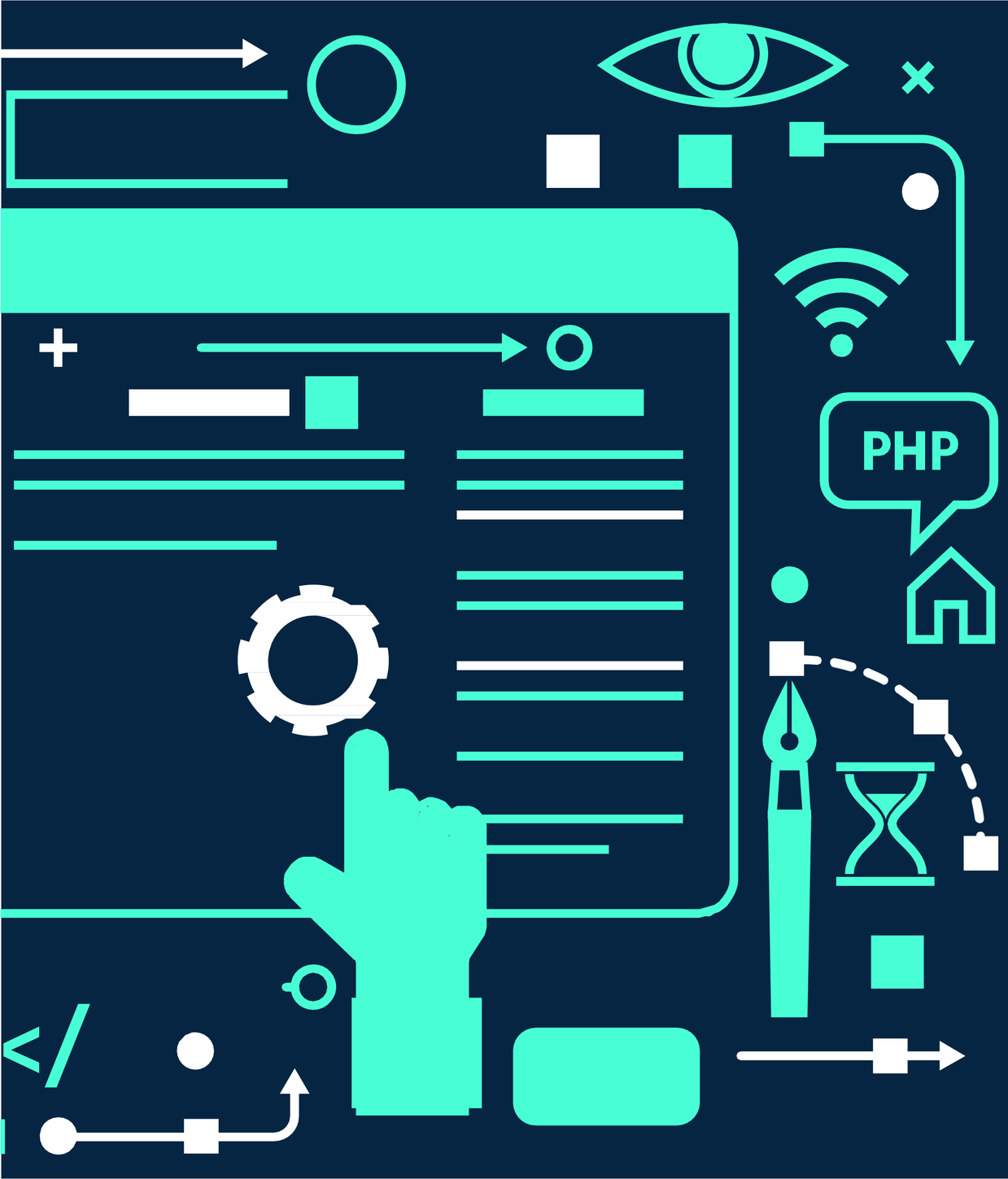
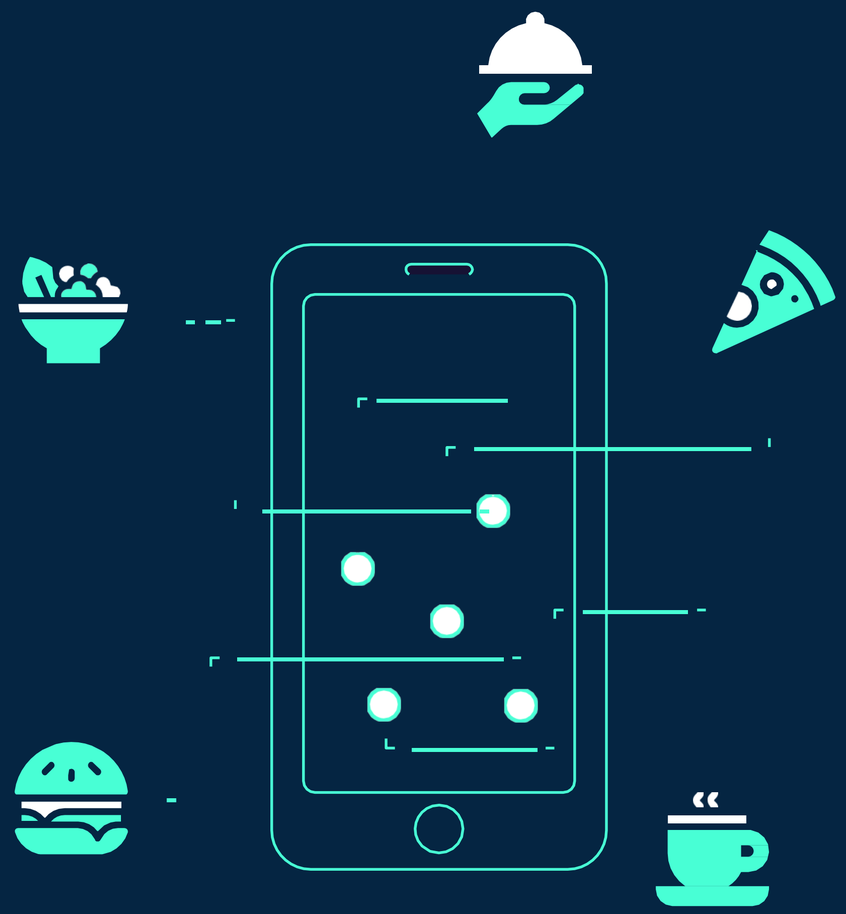


MULTILINGUAL SUMMARIZER

PROBLEM STATEMENT

In the context of information overload, users struggle to efficiently extract meaningful insights from large volumes of text data in multiple Indian languages. Existing summarization tools often lack both accuracy and linguistic diversity, making it difficult to meet the needs of a diverse range of users.



PROJECT SCOPE

- The project aims to develop a multilingual text summarisation website that supports multiple indian languages, such as Hindi , Telugu and English. Provides both **extractive** and **abstractive summarization** options for user flexibility.
- The website integrates hugging face's "csebuetnlp/mt5 multilingual XLSum" model to develop summaries in multiple languages.
- Implement metrics like **Cosine Similarity** and **ROUGE Score** to evaluate the quality of generated summaries and ensure high accuracy.

PROJECT OBJECTIVES

- Develop a robust backend for text processing and summarization.
- Create a user-friendly frontend for easy interaction.
- Ensure high-quality summaries through evaluation metrics (ROUGE, BERT).
- Enhance accessibility for users across different linguistic backgrounds.

APPLICATIONS

- Educational platforms for summarizing course materials.
- News aggregators for concise news summaries.
- Social media analysis to capture trends and sentiments.
- Research tools for literature review and analysis.
- Provide concise summaries of news articles in regional languages, enabling users to quickly consume important information.

CHALLENGES

- Handling linguistic variations and nuances in different languages.
- Ensuring accuracy and relevance in generated summaries.
- Integrating semantic resources effectively with machine learning models.
- Managing computational efficiency for real-time processing.
- Common summarization evaluation metrics may be biased towards English, making it challenging to properly evaluate summary quality in Indian languages.

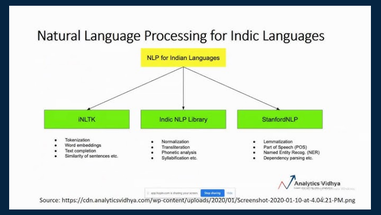
Tools and technologies

01

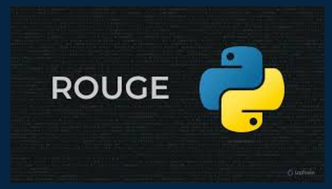
02

03

04

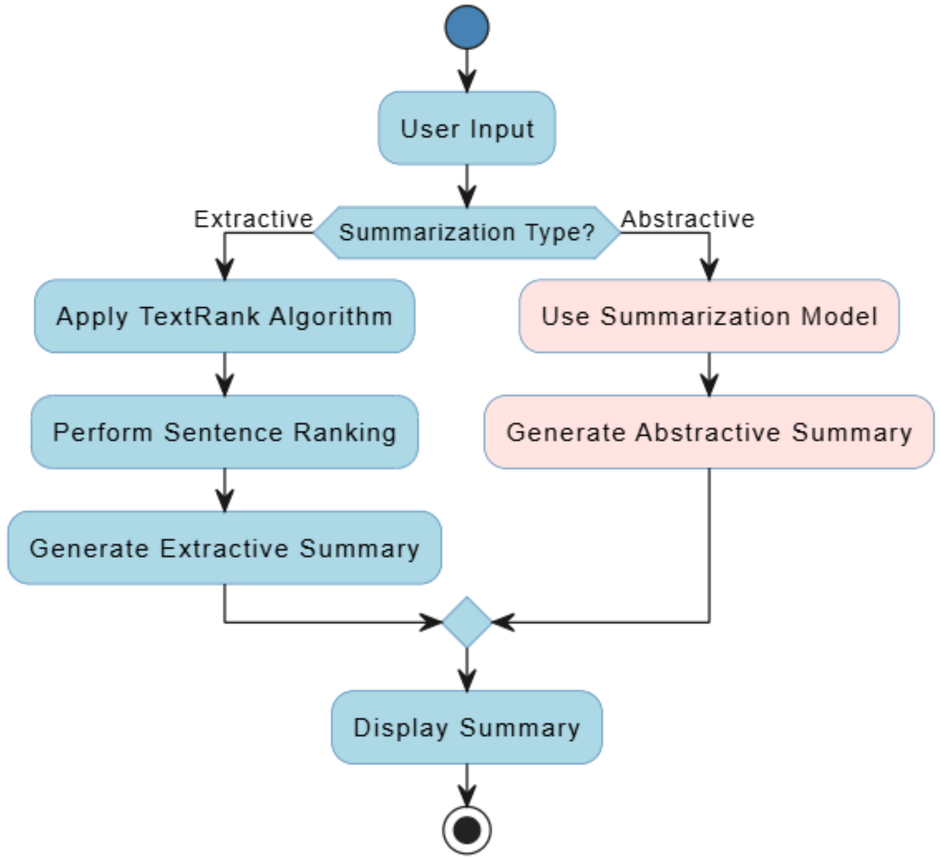


05

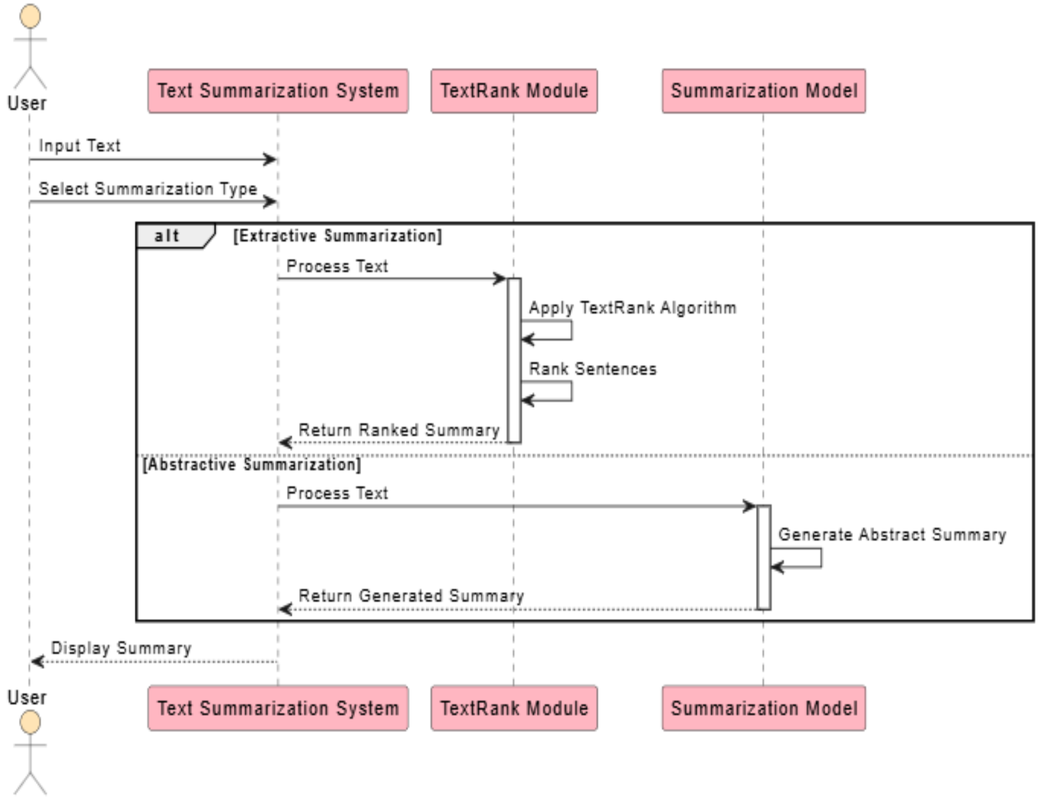


06

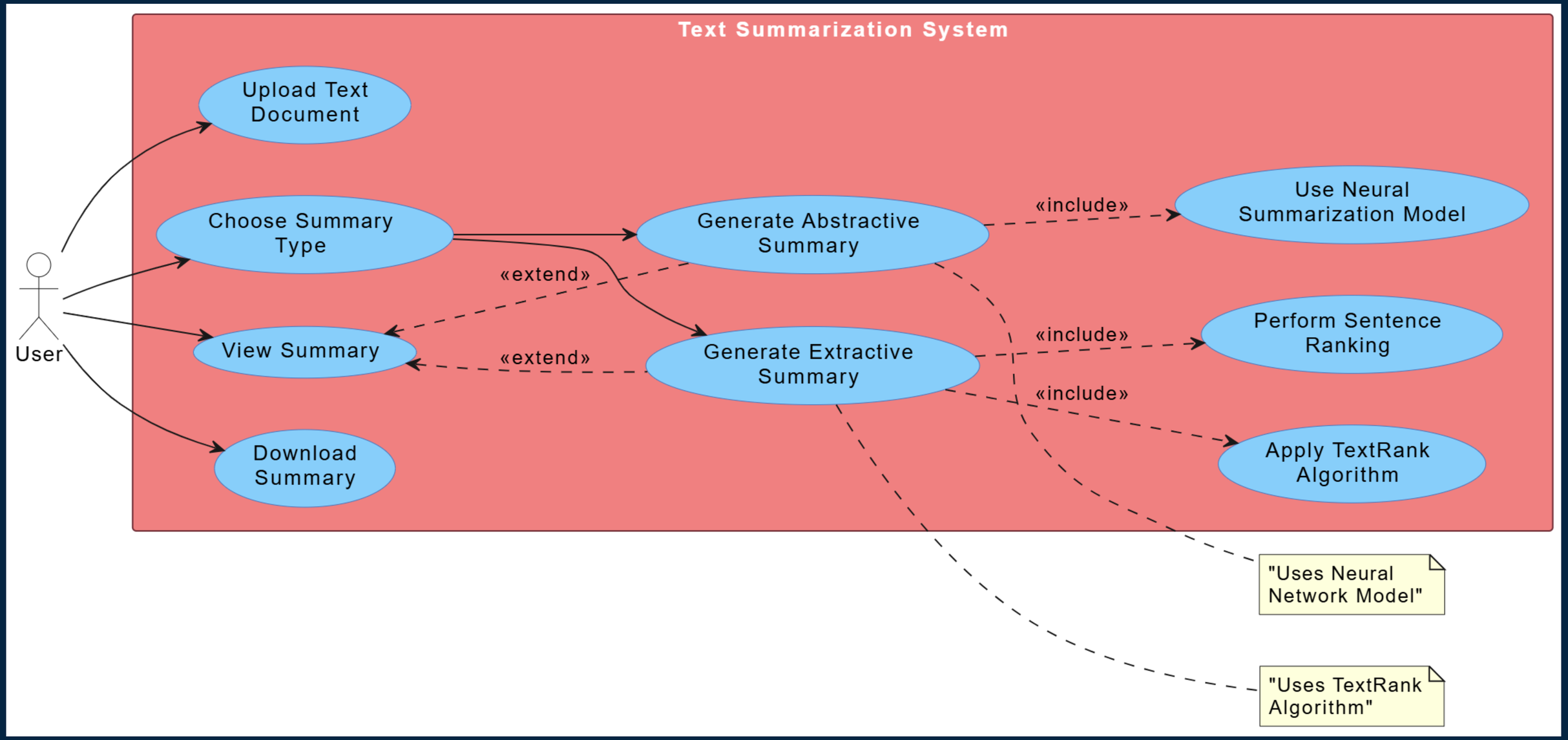
FLOW CHART



SEQUENCE DIAGRAM



USECASE DIAGRAM



PREDICTED RESULTS